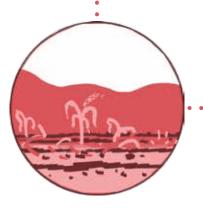## How to Use This Booklet

Data visualization can make information more memorable, more persuasive, facilitate understanding and ultimately motivate action. And within human rights research, it can help investigators and researchers draw a bigger picture from individual human rights abuses by allowing them to identify patterns that may suggest the existence of abusive policies, unlawful orders, negligence, or other forms of culpable action or inaction by decision-makers.

The purpose of this activity is to explore some of the ideas and principles around designing effective data visualization for human rights advocacy.

This activity is broken into a series of five steps each with its own PDF guide. The five steps are:

Step 1: Choose a human rights issue

Step 2: Discuss some kinds of data you might acquire

Step 3: Consider what question are you trying to answer with your data and visualization

Step 4: Choose a chart type for your visualization

Step 5: Consider some data and visualization hazards and ways your charts can be improved

Each step has a corresponding list of options and choices. Read through each step and follow the instructions. Explore the options for each step as you progress.

Step 5

What are some hazards? How could your charts be improved?

Here we will look at how to improve your data visualization as well as some common hazards for data collection, processing, as well as visualization.

Some human rights issue can be difficult to quantify. Listed here are some hazards and tips on collecting data in a way that is accurate and free from bias, and also respects the privacy and dignity of subjects.

Accuracy and clarity are essential to effective visualization, as well as memorability, persuasion, and motivation to action.

When designing your graphics, keep clarity and comprehension in mind and experiment with different chart types that might help your reader understand your story and its key elements.

Distortion

While manipulation of the facts or deception of the reader is usually unintentional in the human rights realm, accidentally misleading visualizations may be deceptive.

Common distortion techniques include: truncated y-axis, representing quantity using area, stretched aspect ratio, inverted axis.

Visual Literacy

Visual literacy, the ability to interpret charts and images, varies widely among and within different populations. Visual literacy is a skill that is learned.

While some audiences may be familiar and comfortable with some chart types, others may be less so. In some cases your audience may be comfortable navigation familiar chart types or novel visualizations. In other cases, it may be best to walk your audience through the stages or features of a given visualization. While for some audiences, a visual or pictorial representation may be preferable to geometric abstraction.

Finally, in some cases it may be best not to use a chart at all and simply present a key data point directly.

The best way to determine the visual literacy of your audience is to test your visuals with representatives of your audience before finalizing your publication.

Spurious Correlation

As the saying goes, "correlation is not causation." Correlation is when two (or more) variables show related trajectories. While this could indicate a causal relationship, there may be other unknown or hidden factors at play.

The careful design of experiments can help rule out some factors, however human rights groups should be extremely cautious and rigorous about making rights claims using correlations.

Non-Zero Baseline

Starting a y-axis at zero has its advantages and disadvantages. Charts where the y-axis does not start at zero can exaggerate the differences between data point values and the steepness of the slope of connecting lines that might otherwise appear rather slight. In some situations this may be deceptive.

However, in other situations slight variation may have enormous impact (for examples, a single degree difference in global warming.) If these small changes are significant to your story, it may make sense to not start the y-axis at zero.

Context should be taken into account, and choosing whether or not to start an axis at zero should be considered with caution.

Uncertainty

One of the great powers of working with statistics is the ability to work with and even quantify uncertainty.

This is not always a familiar concept for the lay views, so visualizing this presents a challenge. Marks on paper imply specificity, so care must be taken when representing a range.

Uncertainty may be expressed as an interval or range of values that reasonably represent the result. Sometimes it is expressed as a range of possible outcomes or paths.

Some ways to represent uncertainty include error bars, fading gradients, violin plots or even dotted lines.

Clear annotation is also key to clarifying what is being represented by the visualization.

Color Value

Color combinations should look good together, but should also be easy to differentiate. Using value (the lightness or darkness of a color) in addition to hue can help make your colors more distinct from each other.

It is easier to perceive approximate differences in length than it is to perceive differences in color, say something is twice as long, vs twice as light. As such, when using colors for categories it is best to group them into visually distinct bins.

Using distinct color values can also help make the chart legible to persons with color blindness.

Color Association

Colors often have cultural associations and this can be an advantage or a disadvantage depending on how colors are used. For instance, election maps in the U.S. use a visual convention of representing blue for Republicans and red for Democrats.

In many charts, the abstract use of color may not evoke cultural associations.

However, when using colors in combination with symbols, or to represent specific countries, populations, or places, one should be mindful about strong cultural color associations, particularly when targeting an international audience.

For example, green may represent luck or "go" in the U.S., while in west Asia it is associated with Islam. The color orange is linked with nationalist, conservative, or liberal political parties in different countries around the world.

For a fascinating glimpse at political color association see
https://en.wikipedia.org/wiki/
Political_colour

Double-Y-Axis

Displaying two different types of data on the same chart make for easy comparison and can save space. However, while it may be tempting to compare two different trends on the same chart, using two different scales on the y-axis can be misleading, particularly if the scales are dramatically different.

Attention is also drawn to the intersection of lines, which may be arbitrary depending on the choice of scales.

If you are convinced that a double-y-axis is the best way to display the data, use very different colors to illustrate the two different data sets. Color code your axes accordingly, to reinforce the data association. If your audience reads from right-to-left, put label your primary dataset on the left y-axis, and your secondary dataset on the right. Avoiding the same chart type for the two data sets can also help clarify the difference. For instance, instead of using two lines, perhaps combine a line for one dataset and a bar chart for the other.

Incomplete Data

Human rights data is almost always incomplete. Minorities and other vulnerable populations may be excluded from data because of deliberate policy, access, resources, or other reasons. The scale or timeline of events may also confound complete data collection.

When using received data, it is important to understand the data collection methods and its limitations.

By using statistical methods like multiple systems estimation, it may still be possible to draw rigorous conclusions from incomplete data.

Violative Collection

Projects collecting data from vulnerable populations should consider risks posed to those populations.

While Universities require a human subjects review, foundations and NGOs do not have the same infrastructure in place.

Even anonymized data can be re-identified when linked with other data sets. This is particularly dangerous where populations are not only marginalized but also criminalized.

For more on data practices for human rights practitioners, read DatNav a guide to using digital data for human rights research, this 2016 report on data ethics, or this handbook from the Responsible Data Forum.

Data Protection

Because of the sensitive nature of working with vulnerable populations, organizations should take care to minimize risk. Data minimization is the practice of limiting the collection of personal data to only that which is directly relevant to the survey. Data anonymization is the process of removing personally identifiable information from data sets.

Critical data is also at risk of a range of protection issues: malware, staff turnover, theft, confiscation, even hardware failure. Care should be taken to encrypt data, to limit access, and to maintain encrypted back-ups in more than one physical location.

Bias

Several types of bias can affect a data visualization at various steps of the process: data collection, processing, analysis, and even visualization.

Selection bias may result from a number of issues including limited access, or victims mistrusting the organization collecting data.

Selection bias particularly affects data related to human rights violations. Selection bias is the selection of individuals, groups, or data for analysis in such a way that it is neither a complete enumeration of all the possible data (like a census) nor a random, scientific sample.

Data Selection

How do you know you're selecting the right data samples to investigate? Is your metric the right one?

It may be tempting to use data that is cleaner or easier to access, but this may also be misleading or more easily taken out of context or may miss the bigger picture.

For instance: focusing on a list of killed human rights defenders may overlook or even downplay tortured individuals, detained, displaced, or silenced by repressive situations.

Appropriate Scale

Audience Analysis

Audience Participation

Simplification

Though truncating the y-axis is generally discouraged, there may be some cases where a series small comparisons may need a truncated axis to be visible. Using appropriate scale in this way can clarify the information, though one should be careful not to deceive the user into thinking the comparisons are more exaggerated than they are.

Knowing your audience can help decide many aspects of your visualization and advocacy. Once you define your target audiences you can start to drill into their motivations: what do they care about? What do they fear? How can you help them take the actions you want? How can you mobilize your allies to reach them?

With target audiences in mind one can start to make decisions about messaging and framing, format and style, the level of simplicity vs complexity, etc. How can you best reach them? What kind of data literacy do they have? What is the best format to reach them? How you can you apply pressure or incentivize them to take appropriate action?

In addition to consent from research subjects, where possible data collection and selection should be conducted in partnership with the subject of its research, for instance, the selection of meaningful indicators.

Once you have developed your data graphics, it's a good idea to test them out, not just with sympathetic staff and allies, but with folks whose profile approximates your target audiences.

Simplification is a matter of removing design elements from a visualization that do not contribute knowledge or insight. Unnecessary design elements could include borders, heavy grid lines, shadows, extraneous colors, gradients, textures, or other decorative elements. These are also referred to as "chartjunk" by Edward Tufte in his book *The Visual Display of Quantitative Information*.

The result of simplification is that it reducing the number of visual elements on a page competing for attention. This reduces cognitive load and can help draw the reader's attention to the most important findings. Combining simplification with emphasis is a powerful way to increase the force and focus of one's message.

Emphasis

In communications for advocacy, time is short and its best to call attention to the main takeaway as quickly and clearly as possible.

One way to do this is creating a visual hierarchy within your information.

Visualization studies also found that the most recognizable visualizations had a clear center of interest. Eye-tracking software determined that the most recognizable visualizations had a single focus, whereas the least recognizable visualizations had less singular patterns.

Visualizations can use emphasis and contrast by differentiating the key figure with color, shape, length, or other visual characteristic.

The order of marks and labels can also create both a visual and logical hierarchy: categories can be ordered alphabetical or by value.

Title and Annotation

Clear titles and text annotation help readers interpret and understand abstract marks and visualizations, and help provide narrative framing.

Titles and annotation also help readers recall the message. Across all textual elements, the title is among the most important. Specific titles reinforcing the visualizations main message often aided recall. Visualizations with non-specific titles can benefit from other explanatory texts such as paragraphs and labels. And visualizations with titles at the top, above the visualization also improved recall.

Titles and annotation should be descriptive and clear, but also concise. Too much text can overwhelm and compete with the visualization. Similarly, font choice and text styling should not compete with the graphics.

Narrative Framing

Abstract shapes and colors in themselves do not carry meaning. Careful annotation can assign meaning, scale, and scope to the representation, but narrative framing can help readers understand the meaning and significance of a given data visualization, and add emotional impact.

One form of narrative framing walks readers through the issue and the graphic, building the story step-by-step.

Stories have a beginning, middle and end. They have actors, evoke emotions and a meaningful climax. Stories guide the audience on a journey.

Data visualizations can be used to walk a reader through an issue in stages. A given data visualization can animate or change over the course of the story, or else different visualizations can be used at different stages of the story.

Concrete Scale

A "concrete scale" is a technique that can be used to visually relate, re-express and compare measures of extreme magnitudes or unfamiliar units.

Concrete scales use familiar and culturally appropriate visual metaphors to depict simple relationships of complex measurements, making it easer for a viewer to comprehend the measure.

Examples of this are: conversion/comparison, containment/nesting, unitization (defining an object as a new unit of measure), and analogy.

For more about concrete scales, see Using Concrete Scales: A Practical Framework for Effective Visual Depiction of Complex Measures.

Redundancy

Data redundancy refers to visually encoding the data in more than one way—for example, including the unit values on a bar chart where the length of the bar also indicates value.

Redundancy helps recall and understanding. Both message and data redundancy make visualizations more memorable. Message redundancy presents the conclusion or message of the visualization in multiple ways to the viewer.

Visual redundancy should be balanced against clutter, and should not detract from the charts emphasis.

Near and Far

This presentation technique balances the macro view with close up of individual data points, for instance combining individual profiles, photos, or testimonies with a statistical view of broader systemic patterns.

Studies have shown that it is difficult for readers to grasp very large numbers outside of their experience, or to empathize with large numbers of victims. The effect, known as "psychic numbing" describes how empathy and willingness to act decreases as the number of victims increases.

The "near and far" approach combines the empathy and personal, human interest angle as context for and of the larger, numeric trends.

This article on the micro-macro view of data visualization goes into more detail.

Visual Grammar

If using many charts through a project, it may make sense to use consistent chart types throughout. While using a variety of forms and formats may make for a visually textured project, avoid asking the reader to compare too many different chart types. This creates additional cognitive load and can be distracting.

Related to consistency of chart types, it is also a good idea to establish a consistent use of color for data elements that appear in a variety of different charts. Establishing a consistent color palette or a grammar of how color is used throughout a project enables the reader to follow the thread of your story through the project with less effort.

Multiples

Instead of trying to cram too much detail into a given chart or map, it may make sense to create multiple versions of the same graphic format, each showing different data or a different state.

Multiples (sometimes called "Small Multiples") are charts or maps using the same scale and axes but displaying different data. Using the same chart size and format allows the user to visually compare similarities and differences. Multiples allow for a more nuanced comparison compared to trying to show everything in one chart.

Tone & Visual Rhetoric

The tone of a graphic is determined by the sum of its details: the style of the typography, the rendering of marks, the voice of the text, the color space, the choice of imagery and image treatment. These add up to creating an overall mood. Often it may be advantageous for a graphic to appear objective, neutral and dispassionate. At other times a more sensational, urgent, or emotive approach may best serve your goals.

Sometimes referred to as "aesthetics," the choice of tone of your graphics should serve both your information as well as your advocacy approach and target audiences, and ideally take cues from your overall communications style and strategy.

Visual Novelty

A powerful way to catch people's attention is the use of visual novelty or user experience. Unexpected visual formats (or even familiar forms put into an unexpected context) can pique an audience's curiosity.

However, this should be balanced by the fact that audiences tend to read best what they read most — that is, chart types that are the most familiar. An audience confronted with an unusual form or experience may be confused and may need to be guided towards interpreting and understanding the presentation.

Visual novelty should remain in the service of your data, your message, and your advocacy objectives. Gratuitous novelty also has the potential make a project easy to dismiss as frivolous.

Additional Resources

The Visual Display of Quantitative Information by Edward Tufte

Two books published by Nancy Duarte offer an excellent primer for visual storytelling: *Slide:ology: The art and science of creating great presentations* and *Resonate: Present Visual Stories that Transform Audiences*

Thinking with Type by Ellen Lupton

About this Booklet